Journal Title Here, 2022, 1–7
doi: DOI HERE
Advance Access Publication Date: Day Month Year
Paper

Mondrian Map: Piet Mondrian's Abstraction for Differential Pathway Analysis

Fuad Al Abir [1] **and Jake Y. Chen** [1, 2, *]

[1]Department of Computer Science, College of Arts and Sciences, The University of Alabama at Birmingham, Birmingham, AL 35233, USA and [2]Systems Pharmacology AI Research Center (SPARC), School of Medicine, The University of Alabama at Birmingham, Birmingham, AL 35233, USA

*To whom correspondence should be addressed.

FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

Abstract

Motivation: Visualizing complex biological networks is a significant challenge, as traditional tools often struggle to represent data clearly and intuitively. Drawing inspiration from Piet Mondrian's abstract art style, we introduce Mondrian Map, a novel visualization tool that transforms the complexity of biological networks into a more organized, meaningful, and aesthetically appealing form.
Results: In our case study of glioma progression, Mondrian Map reveals distinct pathway patterns across multiple patient profiles at different time points. Afterwards, the significance of these distinctive pathways in glioblastoma multiforme development is validated by recent literature. Mondrian Map's visually intuitive representation of complex biological networks enables researchers to easily identify crucial pathways and potential therapeutic targets. Moreover, The potential applications of Mondrian Map extend beyond biology, making it a versatile tool across domains.
Availability and Implementation: The code was implemented in Python version 3.10 and is available through GitHub (github.com/fuad021/mondrian-map). The datasets used in this paper were retrieved from Glioma Longitudinal AnalySiS (GLASS) consortium.
Contact: jakechen@uab.edu
Supplementary information: Supplementary data are available at Bioinformatics online.

1. Introduction

Diving into the intricate world of biological networks is like navigating a labyrinth of interactions where traditional visualization tools often stumble, failing to capture the essence of the data in a clear and intuitive manner. It's a challenge that has left researchers searching for innovative solutions to unlock the secrets hidden within these complex networks. Drawing inspiration from the revolutionary artistry of Piet Mondrian, whose bold strokes and vibrant colors broke the mold of convention, we introduce Mondrian Map — a novel visualization tool poised to revolutionize the way we perceive and understand biological networks. With its fusion of geometric abstraction and vivid primary colors encased in bold borders, Mondrian Map promises to transform complexity into clarity, offering researchers a fresh perspective and a new dimension to explore the intricate tapestry of biological interactions. In Fig. 1, we presented a portrait of Piet Mondrian, one of his art works and a sample Mondrian Map generated by mimicking his art style.

Biological network visualization tools have evolved to address the growing complexity of biomolecular interaction networks. Traditional tools like Cytoscape (Shannon et al., 2003), VisANT (Hu et al., 2004), STRING (Szklarczyk et al., 2023), and Graphia (Freeman et al., 2022) have provided platforms for visualizing various types of biological networks, including

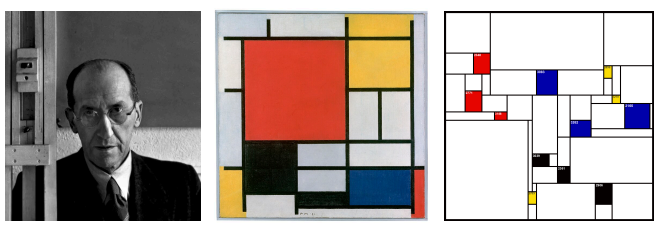

Fig. 1: Mondrian's abstract art inspired network data visualization. Left: Piet Mondrian (1872–1944), photographed by Arnold Newman in 1942. Middle: "Composition with Large Red Plane, Yellow, Black, Grey and Blue" by Mondrian in 1921. Right: Sample Mondrian Map visualization.

© The Author 2022. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

metabolic pathways, protein-protein interactions, and gene regulatory networks. These tools typically represent biological macro-molecules as nodes and their interactions as edges, with annotations represented through different colors, sizes, and distances. However, as networks grow in size, these tools face challenges such as edge-crossings, overlapping nodes, and a general loss of visual clarity due to increased density (Suderman and Hallett, 2007; Albrecht et al., 2010).

GeneTerrain is another visualization framework that uses scattered data interpolation and surface rendering to visualize quantitative attributes of nodes in a network as a continuous surface (You et al., 2010). This technique effectively avoids scalability issues and maintains the topological properties of the original network. However, GeneTerrain's approach can be sensitive to data noise and may present challenges in interpreting the terrain-like visualizations.

While existing visualization tools have provided valuable insights into biological networks, they are often limited by their ability to provide dynamic, time-dependent network analysis. Mondrian Map, with its innovative approach inspired from Piet Mondrian's art style, seeks to address these limitations by offering a novel visualization paradigm that simplifies the complexity of biological networks into more intuitive and visually distinct representations. This approach has the potential to enhance the accessibility and effectiveness of network biology visualizations, facilitating deeper insights into the underlying data. Our aim is to transform the visualization of network biology data into a more organized, meaningful and visually appealing form, where the relationships and key features of the biological entities can be more easily comprehend by the researchers.

Preparing network biology data for Mondrian Map requires several steps. We delve into these steps and the algorithm used to generate Mondrian Map in depth. Subsequently, we present the results along with it's qualitative analysis and validation. Lastly, we conclude the paper by outlining future research directions and discussing the strengths and limitations of the tool.

2. Materials and Methods

Piet Mondrian's abstract art style harnesses the spatial features of fundamental colors, shapes, and lines — possessing an enormous potential for providing an intuitive representation of intricate scientific data. While this approach enhances comprehension, its implementation demands multiple stages of data processing particularly for complex biological networks like pathways. In this section, we delve into the detailed development process of our visualization tool and subsequently address the essential data processing steps required for creating the visualizations.

2.1. Mondrian Map Generation

Characterized by rectangular shapes, primary colors (red, blue, yellow) in bold blocks, and a strict grid structure, Mondrian aimed to achieve a universal and harmonious visual language (Blotkamp, 2001). His compositions often featured horizontal and vertical lines intersecting at right angles, creating a sense of equilibrium and simplicity. In accordance with the salient characteristics and elements of Mondrian's art style, we present a novel algorithm designed for scientific network visualization.

The algorithm necessitates the preprocessing of the input dataset to include entity attributes and relationships. The data can be structured in the form of two tables. The first table encompasses entity attributes such as entity ID, co-ordinate of the central point (x, y), area and color of the entity. And the second table should have entity relationships data such as two entity IDs that shares a relationship and optionally relationship weight for filtering.

Essential attributes such as area, color, and central point co-ordinates are required to represent each entity as a rectangle whereas, the relationships between the entities are illustrated using lines that link the corners of the corresponding rectangles. The definition and significance of these attributes are subject to variability contingent upon the specific use case and the underlying scientific inquiry. For intuitive understanding of the Mondrian Map generation algorithm, we have depicted the key steps in Figure 2. The procedure to create a Mondrian Map is described as follows:

Step 1. Creating Canvas

Initialize an empty canvas featuring a grid system. Then, we plot the central points such that no blocks within the grid system has multiple central points. (Fig. 2 (a))

Step 2. Visualizing Entity Attributes

Utilizing the central points and the designated area, populate an appropriate number of rows and columns of blocks surrounding the point, thereby generating rectangles that are subsequently colored according to specified attributes (Fig. 2 (b)). The approximation of the point and area is deliberately imprecise to maintain adherence to the grid system and achieve aesthetic appeal.

Step 3. Visualizing Entity Relationships

For each entity relationships, identify the two closest vacant corners of the rectangles and establish connections between them using straight line(s), depicted in red, in Fig. 2 (c). In instances where two entities are vertically or horizontally aligned, one

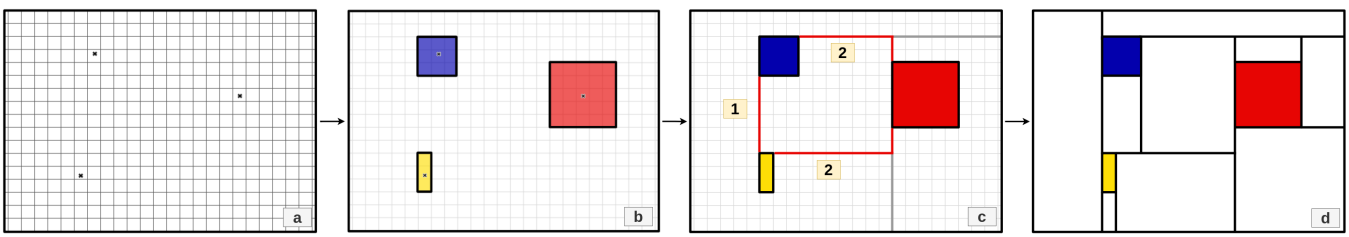

Fig. 2: Implementing Mondrian Map involves 4 key steps: (a) Initiating a grid-based canvas with specified point coordinates, (b) generating rectangles to symbolize entities, (c) illustrating lines to establish relationships between the entities, and (d) populating the canvas with additional lines to adhere to Mondrian's art style.

straight line connection (Type 1) suffices; otherwise, connections are achieved using L-shaped lines (Type 2). For L-shaped lines, we have extended the lines in the original direction of the starting corner, depicted in gray, until the line intersects with the other closest line, rectangle, or border.

Step 4. Mimicking the Art Style

The final step is independent to data and focuses solely on imitating the art style. A vertical and a horizontal line can be drawn from each empty corner of the rectangles, and the smaller line among the two have been rendered to adhere to the cleanliness of the art style. Then, we removed the grid system, and optionally turned all lines to be black for the final output (Fig. 2 (d)).

2.2. Data Preparation

We employed longitudinal gene expression Glioma dataset to demonstrate the effectiveness of Mondrian Map generation. Filtering differentially expressed genesets from gene expression profiles of distinctive patient pools at various time points, we conducted gene-set, network, and pathway analysis to obtain pathway regulatory networks. Subsequently, we generated Mondrian Maps with the networks to find distinctive patterns. This section details the essential data preparation steps required for Mondrian Map generation.

2.2.1. Dataset Introduction

In our study, we utilized the Glioma Longitudinal AnalySiS (GLASS) dataset (The Glioma Longitudinal Analysis (GLASS) Consortium, 2018). The consortium aims to enhance understanding of glioma tumor evolution and identify potential treatment targets and the dataset covers a wide range of patients, from those with highly aggressive gliomas to milder cases. Importantly, it is the only dataset, to our knowledge, that includes gene expression data both before and after interventions across the trajectory of glioma progression. With these unique characteristics, it perfectly aligns with the requirements for effective Mondrian Map visualization for our case study.

2.2.2. Profile selection

We require three distinct patient profiles: aggressive, non-aggressive, and a baseline that falls somewhere in between in terms of aggressiveness. To select these profiles, we followed these steps:

1. Identified patients with at least 3 samples (TP, R1, R2) – only 10 out of 201 patients from the GLASS dataset met this criterion. Here, TP represents the Tumor Primary, while R1 and R2 denote the 1st and 2nd recurrence of the tumor, respectively.

2. From the 10 filtered patients based on clinical data, we chose 2 patients representing the most aggressive and non-aggressive glioma. Patient with ID 0279 had an aggressive course, diagnosed with Glioblastoma (grade IV) at the age of 34, surviving 11 months despite of multiple surgeries. In contrast, Patient 0027, diagnosed with Astrocytoma (grade II) at 30, later gradually developed Glioblastoma (grade IV), surviving 80 months post-diagnosis with distinct surgical interventions.

3. Lastly, we established a baseline profile by averaging gene expression values from five moderately aggressive patients who survived between 37 to 75 months. These patient profiles will enable us to conduct a thorough comparative analysis of Mondrian Maps, aiming to uncover intriguing patterns.

2.2.3. Differentially Expressed Gene Sets

Our study examines three distinct patient profiles: aggressive, baseline, and non-aggressive and each patient profile has three gene expression profiles: TP, R1, and R2. To refine our dataset, we applied a threshold of 0.001 to remove genes with nearly zero expression levels across all nine profiles. This narrowed down our dataset to of 11,075 genes for further analyses.

Subsequently, we computed the Fold Change (FC) values for R1 and R2 relative to TP, represented as R1/TP and R2/TP. Following this, we applied thresholds for up-regulation ($>= 1.5$) and down-regulation ($<= 0.5$) to identify significant changes in gene expression. This process yielded a total of six differentially expressed gene sets.

For the aggressive patient, we identified 741 up-regulated and 922 down-regulated genes for the 1st recurrence (R1) respective to the primary tumor (TP). Whereas for the 2nd recurrence (R2), 980 genes exhibited up-regulation, while 1100 genes were down-regulated. Similarly, for the non-aggressive and baseline profiles, we identified varying numbers of differentially expressed genes (DEGs). Number of genes for the differentially expressed gene sets can be found in Supplementary Table 1.

2.2.4. Retrieving Pathway Information for DEGs

Utilizing the six DEGs from section 2.2.3, we conducted gene-set, network, and pathway analysis (GNPA) to unravel the interconnected functions of genes, providing a comprehensive understanding of biological processes and their implications for specific conditions or diseases. To achieve this, we used the Pathway, Annotated-list, and Gene-signature Electronic Repository (PAGER) API (Yue et al., 2015) to obtain pathway data. A PAG (Pathway, Annotated-list, and Gene-signature) denotes a set of genes sharing common features, including pathways, biological processes, co-differential expression under specific conditions, or common downstream targets. PAGER serves as an online systems biology tool, facilitating the construction and visualization of PAG-to-PAG regulatory networks derived from multiple PAGs. This tool provides complementary biological insights, extending beyond standard gene set analysis or individual gene network analysis (Chen et al., 2006).

For uniformity, we exclusively queried WikiPathways as the source and considered only the statistically significant PAGs ($p - value < 0.05$). We have used PAGs and pathways interchangeably in this paper, acknowledging that pathways are a subset of PAGs and as we only used WikiPathways as the source. Therefore, we only retrieved pathways from the PAGER API as a result. Subsequently, we retrieved the m-type PAG-to-PAG co-membership network using the PAG IDs. The m-type relationship networks connect pairs of PAGs with a significant number of shared genes (Yue et al., 2015). A summary of PAGs and PAG-to-PAG relationships is presented in Supplementary Table 2.

2.2.5. Differential Pathway Analysis

We have retrieved all the genes associated in those pathways and their RP-scores using PAGER API to calculate weighted FC for each PAGs. The RP-score, stands for rank prioritization-score, is used to rank genes within the PAG (Yue et al., 2015). It is calculated and assigned to each gene in the PAG using the Eq. 1.

$$RP-Score = e^{k \times ln(\sum_{p,q \in PAG} conf(p,q)) - ln(\sum_{p,q \in PAG} N(p,q))} \quad (1)$$

where p and q as gene indexes from the PAG, with a constant $k(=2)$, the term $conf(p,q)$ represents the interaction confidence score assigned by HAPPI-2 (Chen et al., 2017), ranging between 0 and 1. Additionally, $N(p,q)$ takes on a value of 1 if gene p interacts with q, or 0, otherwise.

With the fold change (FC) values calculated in section 2.2.3 and the RP-score as weights (W), we have calculated weighted fold change (wFC) for each PAGs using Eq. 2.

$$wFC = \frac{\sum_{i=1}^{n} W_i * FC_i}{\sum_{i=1}^{n} W_i} \quad (2)$$

where n is the number of genes in the PAG, W_i is the RP-score of $Gene_i$ and FC_i is the fold change value of $Gene_i$. For the downstream analysis, we consider a pathway to be up-regulated if $wFC \geq 1.25$ and down-regulated if $wFC \leq 0.75$.

2.2.6. Pathway Representation

The definition of the visualization attributes can vary depending on the specific use case. In our cause, we depict each PAG as a rectangle and the m-type relationships using lines that connect the corners of these rectangles. The size of the rectangles corresponds to the absolute log2 of wFC, and we position the central point of each PAG on the canvas with force-directed layout (Fruchterman and Reingold, 1991) using the networkX python library (Hagberg et al., 2008).

Designing the layout, we have developed guardrails to ensure no overlaps between the rectangles and minimizing the intersection between the lines to make the visualizations clean. In the layout, the locations of each PAGs are preserved so that the PAGs remain in the same position throughout the analysis. This is essential for comparative analysis between two Mondrian Maps.

All the PAGs obtained using the PAGER API show significance in terms of their p-value. However, some of them can be insignificant when considering the adjusted p-value. In such cases, these PAGs are depicted as black rectangles. Conversely, those PAGs deemed significant based on the adjusted p-value are highlighted: red for up-regulated, blue for down-regulated, and yellow for other cases. The number of different types of differentially expressed pathways is presented in Supplementary Table 3.

We also color-coded the relationship lines, inspired from Mondrian's New York series (Mondrian, 1942–1944) as varying line colors can convey additional visual cues. Lines are red if it connects two up-regulated PAGs and blue if both of the PAGs are down-regulated. The rest relationship lines to be yellow whereas the gray lines has no associated relationships.

In our retrieved data from the PAGER API, the PAG IDs have the common format of WAG00xxxx. Therefore, we have annotated the last 4 digits of PAG ID at the top-left corner of each rectangle for further analysis and validation. At this point, we have all the data retrieved and calculated for our visualization.

2.3. Qualitative Validation

To validate the patterns we identified by analyzing Mondrian Maps, we used a language model-based search engine called Perplexity AI (Perplexity AI, 2022). By querying the PAGER Web App (Yue et al., 2022) with specific PAG IDs, we obtained details about the pathways we were interested in. With this pathway information in hand, we then queried the search engine to explore any connections between the disease and these pathways. The search engine promptly provided us with references to recent literature indicating any relationships found. This process allowed us to delve into the relationships between the observed PAGs and GBM, bolstering the credibility of our findings.

3. Results

After thoroughly describing the Mondrian Map implementation in section 2.1 and outlining the data processing steps in section 2.2, this section delves into a comprehensive analysis of the visualizations. We highlight distinctions and similarities among various profiles at different time points, followed by biological validations of our proposed method.

3.1. Mondrian Map Analysis

To reiterate, in this case study, we have baseline, aggressive, and non-aggressive glioma patient profiles where each of the profiles have 3 gene expression samples in different time points, namely TP (tumor primary), R1 (1st recurrence) and R2 (2nd recurrence). As we are using fold change to define the area for each of the rectangles, we can generate two Mondrian Maps from each of the patients such as R1/TP and R2/TP, resulting six in total. We have presented all the generated Mondrian Maps in Fig. 3.

Creating these Mondrian Maps, we have filtered the top 10 most significant PAGs in terms of adjusted p-value. In the first row of Fig. 3, we have assembled the R1/TP Mondrian Maps and R2/TPs in the second. The patient profiles are sorted column-wise i.e., baseline, aggressive and non-aggressive profiles, respectively.

In the baseline profile, all the top 10 PAGs are red, i.e., up regulated (Fig. 3 A & D). We find the area of PAG ID 3032, 3212, 2628, 3066 and 2659 are increasing in R2/TP compared to R1/TP. That means the absolute log2 of wFC for the respective PAGs in R2/TP are greater than that of R1/TP. The significance of increasing in area is a PAG is regulating in greater fold comparatively. We can find similar area increment in R2/TP compared to R1/TP for other profiles as well.

In Fig. 3 B & E, the not differentially expressed PAGs (yellow) in aggressive profile exhibit a smaller area coverage due to their lower log2 of wFC compared to the up-regulated PAGs in the baseline. PAGs with IDs 2659, 2393, 2805, 3294, and 3196 show an increase in area in R2/TP compared to R1/TP. Notably, two down-regulated (blue) PAGs in R1/TP are absent in the R2/TP Mondrian Map, while two new up-regulated (red) PAGs emerge in R2/TP. Additionally, all the m-type relationships between the PAGs in R1/TP also persisted in R2/TP.

Notably, PAG with ID 3212 stands out in the top 10 PAGs for all six Mondrian Maps. In the aggressive R1/TP and R2/TP, it appears as a small, non-differentially expressed yellow rectangle whereas in the non-aggressive Mondrian Maps, the area increased, but it remains yellow. However, in the baseline Mondrian Maps, PAG 3212 is distinctly up-regulated, marked by the large red rectangles. Additionally, this PAG forms two m-type relationships with other PAGs in baseline and aggressive profiles, underscoring its significance for further analysis.

Furthermore, all the PAGs in non-aggressive R1/TP Mondrian Map are yellow, i.e., not differentially expressed. However, in R2/TP Mondrian Map, PAG ID 2133, 2628, 2876, 2483 and 2783 are up regulated and the increment in area is in many folds. PAGs in lower-right portion of R1/TP and R2/TP for non-aggressive Mondrian Maps (Fig. 3 C & F) are labeled with their IDs as they

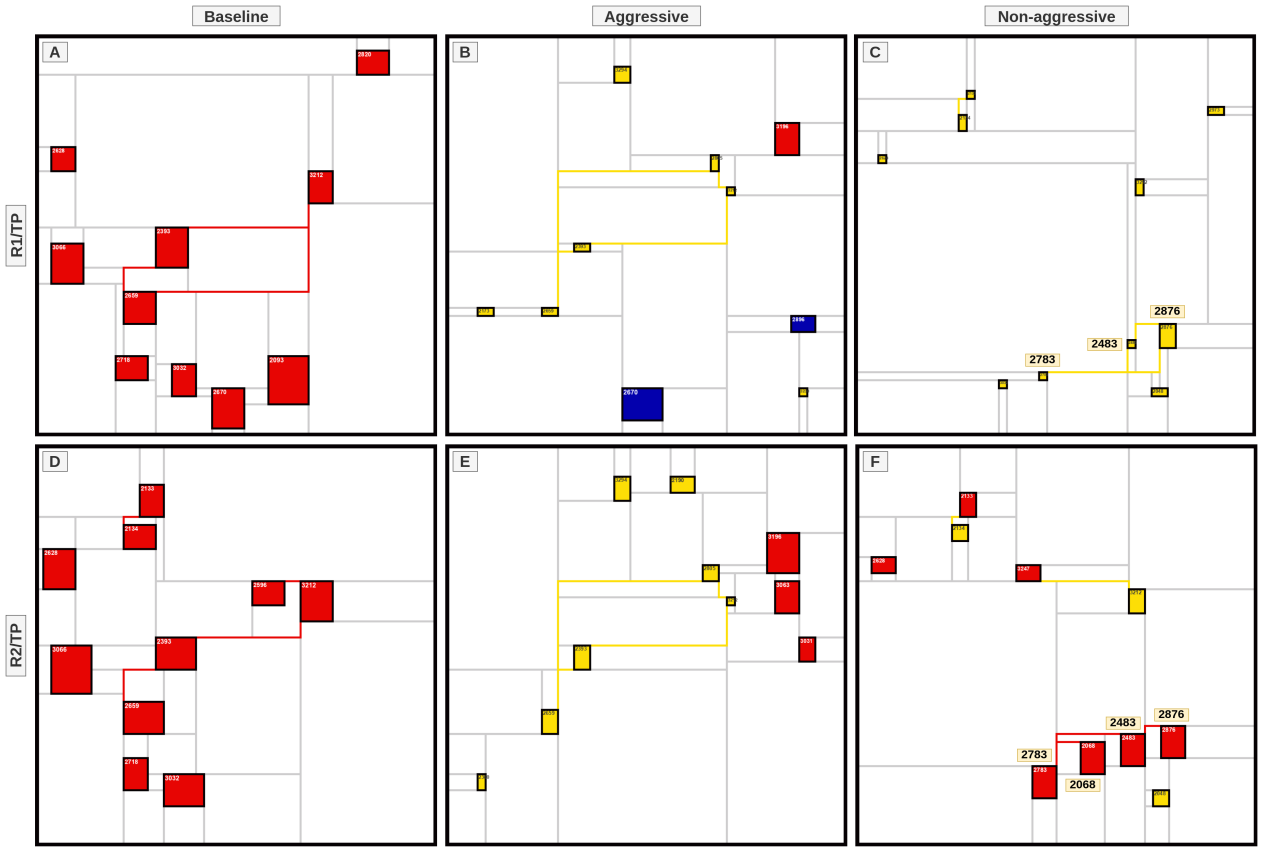

Fig. 3: Mondrian Maps for three patient profiles and two recurrence respective to primary tumor diagnosis. Profiles have been sorted in columns whereas the rows have R1/TP and R2/TP, respectively. PAGs from the bottom right corner on non-aggressive patient are explicitly labeled with PAG IDs for further explanation.

show most distinctive features. Therefore, these PAGs demand further discussions to understand the distinctive states of glioma progression and we have presented detailed analysis by validating our claims with recent biological literature in section 3.2. This analysis can be done similarly for other distinctive PAGs as well.

3.2. Biological Data Validation

To validate the visual cues from the Mondrian Maps, we have focused on the lower-right portion of non-aggressive patient profile depicted in Fig. 3 C & F. We can see PAG 2783, 2483 and 2876 has become yellow to red and increased in size as the tumor progresses and PAG 2068 in R2/TP has PAG-to-PAG relations with the other PAGs. Searching the PAGER web app with the IDs, the titles of the pathways are respectively Cilium Assembly, Assembly of the primary cilium, Ciliary landscape and Mitotic G2-G2/M phases. From the pathway titles, it is very intuitive that cilia may have a significant role in glioma progression. In the recent literature, we found multiple evidences that support our claim.

Primary cilia are microtubule-based organelles that protrude from the apical surface of most mammalian cells. In the human brain, primary cilia are essential for maintaining the proper flow of cerebrospinal fluid (CSF) and facilitating communication between cells (Park et al., 2019; Hasenpusch-Theil and Theil, 2021). They are typically present during the G0 or G1-S/G2 phases of the cell cycle and play a crucial role in cell signaling and regulation (Plotnikova et al., 2009).

Emerging evidence suggests that abnormalities in primary cilia are involved in GBM progression (Sarkisian and Semple-Rowland, 2019). In GBM biopsies, many tumor cells lack cilia, indicating that the absence or dysfunction of these organelles may contribute to tumorigenesis (Deleyrolle and Sarkisian, 2023). Dysregulated cilia can lead to impaired CSF flow, creating a microenvironment that promotes tumor growth and invasion (Banizs et al., 2005). Additionally, alterations in the ciliary landscape, such as defects in cilium assembly or disassembly, can disrupt critical signaling pathways that regulate cell proliferation and differentiation (Wang et al., 2021).

The mitotic G2-G2/M phases are also implicated in glioma progression (Vitovcova et al., 2020). GBM cells exhibit alterations in cell cycle regulation, particularly in the G2-M phase, which is crucial for cell division. The G2-M checkpoint ensures that cells do not enter mitosis with damaged DNA or incomplete replication. Abnormalities in this checkpoint can lead to uncontrolled cell proliferation, a hallmark of cancer. Targeting the G2-M phase and its regulators, such as cyclin-dependent kinases, Aurora kinases, and surviving, has shown promise as a therapeutic strategy for GBM (Castro-Gamero et al., 2018).

The interconnectedness of these pathways highlights the complexity of GBM biology. Alterations in primary cilia and the ciliary landscape can affect cell cycle regulation, while

abnormalities in the G2-M phase can impact cilia function (Plotnikova et al., 2009). Furthermore, there are significant interest among the researchers to find the role of cilia in glioma as we have found more that 700 studies since 2023 in google scholar when we searched with two terms, "Cilia" and "Glioma". Understanding the interplay between these processes is crucial for developing effective therapies targeting GBM.

Backed by the recent literature, it is established that the mitotic G2-G2/M phases and cilia play significant roles in glioma progression and the development of GBM. Using Mondrian Map, we can easily find such crucial pathways which can be potential targets for therapeutic intervention. Further research into the molecular mechanisms underlying these processes may lead to the development of novel treatment strategies for glioma.

4. Discussion

Mondrian Map's key strength lies in its ability to simplify and visually present intricate biological networks. Parsing through tables with extensive numerical values is challenging for humans. By providing a visually intuitive understanding of high-level information, Mondrian Map facilitates comparative analysis across different stages and profiles, guiding the identification of novel therapeutic targets. This section delves into the strengths and limitations of Mondrian Map while shedding light on potential avenues for further development.

4.1. Strengths and Limitations

The Mondrian Map visualization tool offers a unique approach to presenting complex biological networks in an intuitive and visually appealing manner. Its strength lies in its ability to distill intricate data into a simplified representation, making it easier for researchers to grasp high-level ideas and patterns at a glance. The tool's aesthetically pleasing design not only enhances the user experience but also facilitates effective communication of findings to a broader audience.

Furthermore, Mondrian Map excels in enabling downstream analysis by allowing users to zoom in on sub-networks. This proves invaluable when investigating specific interactions or pathways within larger systems, providing researchers with a focused view and deeper insights into the underlying biological processes.

Despite its strengths, Mondrian Map faces certain limitations that should be acknowledged. When dealing with highly complex networks characterized by dense interconnections, the visualization can become overwhelming and challenging to interpret. Currently, the tool is suitable for networks with node degrees less than or equal to four at most (one relationship per corner of the rectangles), as higher degrees can lead to visual clutter and obscure the underlying relationships. Moreover, the number of pathways and relationships are controllable. Therefore, complex networks can be mitigated by having larger canvas and prioritized pathways and relationships to visualize.

Additionally, the process of finding an optimal layout without any overlaps for complex datasets can be time-consuming and computationally intensive. In some cases, the layout algorithm may even fail to converge, resulting in unsatisfactory visualizations due to issues such as excessive overlap between the rectangle and/or intersecting connecting lines. These challenges can arise from the inherent complexity of the data, hindering the tool's ability to provide a clear and interpretable representation.

There might be concerns regarding how the rectangles are generated in Mondrian MAP, as outlined in Section 2.1, Step 2. It is important to note that while the area of each entity may not precisely equal to the given value, it is approximated to maintain minimal visualizations of small blocks and to create regular shapes with aesthetically pleasing aspect ratio. To address potential imprecision, optimizing the grid size in the initial step of Mondrian Map Generation (Section 2.1) can be done.

Moreover, the effectiveness of Mondrian Map's visualization is influenced by several hyperparameters, including area scalar, grid size, and guardrails. Tuning these parameters based on the specific use case is crucial to achieving optimal results, as different datasets may require different parameter settings to ensure clarity and readability. Finally, as we made the software freely available, more case studies will be added in GitHub and future application notes.

4.2. Future work and Conclusions

While the current implementation of Mondrian Map effectively utilizes rectangles, their area, colors, and lines to depict relationships between the entities, several future enhancements can further enrich the visual experience. Drawing inspiration from Mondrian's iconic works, such as Composition II, Composition no. 10, and the Boogie Woogie series, the tool can incorporate features like line thickness to represent connection strength and different line types for added dimensionality. Additionally, allowing for horizontal or vertical rectangle orientations can enable the representation of additional state data, further expanding the tool's expressive capabilities.

The potential applications of Mondrian Map extend far beyond the realm of biology, making it a highly versatile and adaptable tool across numerous domains. From demographic data visualization, where block sizes represent population sizes, colors denote demographic groups, and positions signify geographic regions, to economic indicators, where GDP, economic sectors, and geographic locations are mapped using the same visual language, Mondrian Map offers a fresh and intuitive approach. Other promising areas include web traffic analysis, sales data representation, climate data mapping, healthcare and educational statistics, energy consumption patterns, transportation network visualization, and social network analysis to name a few.

While Mondrian Map excels in visualizing temporal networks, its current implementation may face limitations with higher degree nodes. To address this, a potential solution could involve introducing midpoints along the sides of the rectangles, effectively creating eight distinct points for relationship lines, thereby accommodating more connections. Furthermore, the color scheme can be enhanced by introducing more colors, and incorporating transparency or gradient effects akin to heatmaps to represent intensity levels. These adaptations would not only improve the tool's ability to handle intricate networks but also provide additional layers of information and visual clarity, further solidifying Mondrian Map's position as a powerful and versatile visualization tool across diverse domains.

In conclusion, our proposed visualization technique emerges as a transformative tool for unraveling the intricacies of complex biological networks. Mondrian Map transcends the limitations of numerically rich tables with the state of the art visualization techniques, providing an intuitive understanding that facilitates comparative analyses across multiple stages and patient profiles.

The visual insights generated by Mondrian Map not only enhance the interpretability of intricate interactions but also pave the way for the identification of novel therapeutic targets. Our study lays the groundwork for future advancements and refinements in this promising visualization technique within the realm of bioinformatics and network biology and beyond.

5. Competing interests

No competing interest is declared.

6. Author contributions statement

J.C. conceived the idea, designed the approach, and supervised the implementation process, F.A.A. developed software, conducted the experiments, and analyzed the results. F.A.A. drafted the manuscript and both authors proofread it.

7. Acknowledgments

J.C. acknowledges the support of NIH grant awards U54OD036472, U54DK137307, 1OT2OD032742, and R01HL150078.

References

Mario Albrecht, Andreas Kerren, Karsten Klein, Oliver Kohlbacher, Petra Mutzel, Wolfgang Paul, Falk Schreiber, and Michael Wybrow. On open problems in biological network visualization. In *Graph Drawing: 17th International Symposium, GD 2009, Chicago, IL, USA, September 22-25, 2009. Revised Papers 17*, pages 256–267. Springer, 2010.

Boglarka Banizs, Martin M Pike, C Leigh Millican, William B Ferguson, Peter Komlosi, James Sheetz, Phillip D Bell, Erik M Schwiebert, and Bradley K Yoder. Dysfunctional cilia lead to altered ependyma and choroid plexus function, and result in the formation of hydrocephalus. *Development*, 2005.

Carel Blotkamp. *Mondrian: The art of destruction*. Reaktion Books, 2001.

Angel Mauricio Castro-Gamero, Julia Alejandra Pezuk, María Sol Brassesco, and Luiz Gonzaga Tone. G2/m inhibitors as pharmacotherapeutic opportunities for glioblastoma: the old, the new, and the future. *Cancer biology & medicine*, 15(4):354, 2018.

Jake Y Chen, Changyu Shen, Zhong Yan, Dawn PG Brown, and Mu Wang. A systems biology case study of ovarian cancer drug resistance. In *Computational Systems Bioinformatics*, pages 389–398. World Scientific, 2006.

Jake Y Chen, Ragini Pandey, and Thanh M Nguyen. Happi-2: a comprehensive and high-quality map of human annotated and predicted protein interactions. *BMC genomics*, 18:1–16, 2017.

Loic P Deleyrolle and Matthew R Sarkisian. Cilia at the crossroads of tumor treating fields and chemotherapy. *Developmental Neuroscience*, 45(3):139–146, 2023.

Tom C Freeman, Sebastian Horsewell, Anirudh Patir, Josh Harling-Lee, Tim Regan, Barbara B Shih, James Prendergast, David A Hume, and Tim Angus. Graphia: A platform for the graph-based visualisation and analysis of high dimensional data. *PLoS computational biology*, 18(7):e1010310, 2022.

Thomas MJ Fruchterman and Edward M Reingold. Graph drawing by force-directed placement. *Software: Practice and experience*, 21(11):1129–1164, 1991.

Aric Hagberg, Pieter Swart, and Daniel S Chult. Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

Kerstin Hasenpusch-Theil and Thomas Theil. The multifaceted roles of primary cilia in the development of the cerebral cortex. *Frontiers in Cell and Developmental Biology*, 9:630161, 2021.

Zhenjun Hu, Joseph Mellor, Jie Wu, and Charles DeLisi. Visant: an online visualization and analysis tool for biological interaction data. *BMC bioinformatics*, 5:1–8, 2004.

Piet Mondrian. New york series, 1942–1944.

Sang Min Park, Hee Jin Jang, and Jeong Ho Lee. Roles of primary cilia in the developing brain. *Frontiers in cellular neuroscience*, 13:218, 2019.

Perplexity AI. Perplexity ai. https://www.perplexity.ai/, 2022. Accessed: April 1, 2024.

Olga V Plotnikova, Elena N Pugacheva, and Erica A Golemis. Primary cilia and the cell cycle. In *Methods in cell biology*, volume 94, pages 137–160. Elsevier, 2009.

Matthew R Sarkisian and Susan L Semple-Rowland. Emerging roles of primary cilia in glioma. *Frontiers in Cellular Neuroscience*, 13:55, 2019.

Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S Baliga, Jonathan T Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, 13(11):2498–2504, 2003.

Matthew Suderman and Michael Hallett. Tools for visually exploring biological networks. *Bioinformatics*, 23(20):2651–2659, 2007.

Damian Szklarczyk, Rebecca Kirsch, Mikaela Koutrouli, Katerina Nastou, Farrokh Mehryary, Radja Hachilif, Annika L Gable, Tao Fang, Nadezhda T Doncheva, Sampo Pyysalo, et al. The string database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic acids research*, 51(D1):D638–D646, 2023.

The Glioma Longitudinal Analysis (GLASS) Consortium. Glioma through the looking glass: molecular evolution of diffuse gliomas and the glioma longitudinal analysis consortium. *Neuro-oncology*, 20(7):873–884, 2018.

Barbora Vitovcova, Veronika Skarkova, Kamil Rudolf, and Emil Rudolf. Biology of glioblastoma multiforme—exploration of mitotic catastrophe as a potential treatment modality. *International journal of molecular sciences*, 21(15):5324, 2020.

Bo Wang, Zheyong Liang, and Peijun Liu. Functional aspects of primary cilium in signaling, assembly and microenvironment in cancer. *Journal of Cellular Physiology*, 236(5):3207–3219, 2021.

Qian You, Shiaofen Fang, and Jake Yue Chen. Gene terrain: visual exploration of differential gene expression profiles organized in native biomolecular interaction networks. *Information Visualization*, 9(1):1–12, 2010.

Zongliang Yue, Madhura M Kshirsagar, Thanh Nguyen, Chayaporn Suphavilai, Michael T Neylon, Liugen Zhu, Timothy Ratliff, and Jake Y Chen. Pager: constructing pags and new pag–pag relationships for network biology. *Bioinformatics*, 31 (12):i250–i257, 2015.

Zongliang Yue, Radomir Slominski, Samuel Bharti, and Jake Y Chen. Pager web app: An interactive, online gene set and network interpretation tool for functional genomics. *Frontiers in Genetics*, 13:820361, 2022.